Ensemble Learning: Building Smarter Models through Collaboration

24-11-2025

Alejandro Morales Hernández
Machine Learning Group
Université Libre de Bruxelles, Belgium

* Supported by INNOVIRIS through the TORRES project
https://torresml6.wordpress.com/

1

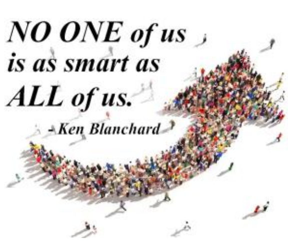
Outline

2

2

Ensembles

- **Increase accuracy:** Combining models usually performs better than using each model individually.
- **Increase robustness:** Ensembles reduce the risk of bad predictions.
- **Tackling complex problems:** Some problems are simply too complex for one model to handle alone.



3

3

The magician (magic)

The bandit (sneaky)

magic + creativity + strength + sneaky = ***Success!***

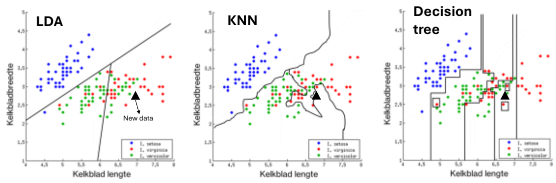
How can an ensemble of ML models (the adventurers) solve a complex problem (a mystery) better than a single model (a hero)?

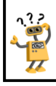
4

4

Base classifiers

Input features → ML model → Output label





It is Virginica

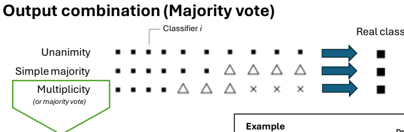
It is Versicolor

It is Versicolor

5

5

Output combination (Majority vote)



$$P(k) = \sum_{i=1}^{L} I(s_i, k)\ ; \quad I(a,b) = \begin{cases} 1; & a = b \\ 0; & a.g. \end{cases}$$

Classifiers

Class k

Class given by classifier i

$$k^* = \underset{k=1\cdots n}{\arg\max}\ P(k)$$

Decision profile

Final class

Example (*Majority vote*):

	Class
Classifier 1	■
Classifier 2	■
Classifier 3	■
Classifier 4	■
Classifier 5	Δ
Classifier 6	Δ
Classifier 7	Δ
Classifier 8	X
Classifier 9	X
Classifier 10	X

Decision profile

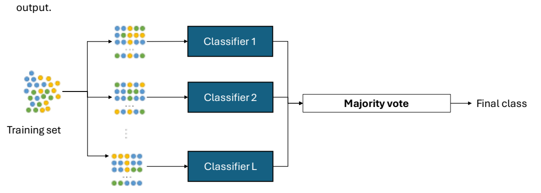
	Class		
	■	Δ	X
Classifier 1	1	0	0
Classifier 2	1	0	0
Classifier 3	1	0	0
Classifier 4	1	0	0
Classifier 5	0	1	0
Classifier 6	0	1	0
Classifier 7	0	1	0
Classifier 8	0	0	1
Classifier 9	0	0	1
Classifier 10	0	0	1

6

6

Ensembles' approaches – "*Bagging*"

- The ensemble contains classifiers* trained with bootstrap-replica's extracted from the "training set".
- Use "**majority vote**" to combine the output of the classifiers.
- The classifier should be instable (e.g. *Decision Tree*): small changes in the "training set" have a large impact in the output.



* The same base classifier

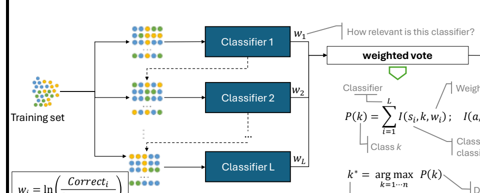
7

7

Ensembles' approaches – "*Boosting*"

* The same base classifier

- The ensemble is built by adding a classifier* each time.
- **The new classifier learns** from a dataset that is specifically designed to **reduce the errors made by the previous classifier**.
- Use "**weighted vote**" to combine the output of the classifiers.



$$w_i = \ln\left(\frac{Correct_i}{Incorrectt_i}\right)$$

Classifier

Weight for classifier i

$$P(k) = \sum_{i=1}^{L} I(s_i, k, w_i)\ ; \quad I(a,b,w) = \begin{cases} w; & a = b \\ 0; & a.g. \end{cases}$$

Class k

Class given by classifier i

$$k^* = \underset{k=1\cdots n}{\arg\max}\ P(k)$$

Decision profile

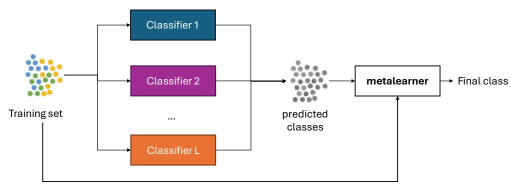
Final class

8

8

Ensembles' approaches – "*Stacking*"

- "Stacking" is mostly used to combine different classifiers.
- A "**metalearner**" uses the classes predicted by the classifiers as input features and the target output is the same as in the original "training set".



9

9

Applications

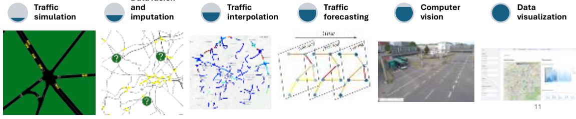
10

Traffic processing foR uRban EnvironmentS (TORRES)

https://torresml6.wordpress.com/

- Provide **authorities** and **cities** with the means to better understand and **quantify** the **impact of their policies** on traffic and mobility, which directly relate to **citizen's quality of life** and **safety**.
- Using AI and machine learning to allow authorities and cities to make smarter data-driven decisions.

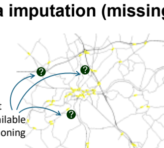
Traffic simulation | Data fusion and imputation | Traffic interpolation | Traffic forecasting | Computer vision | Data visualization



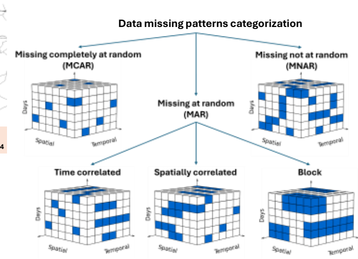
11

11

Data imputation (missing patterns)



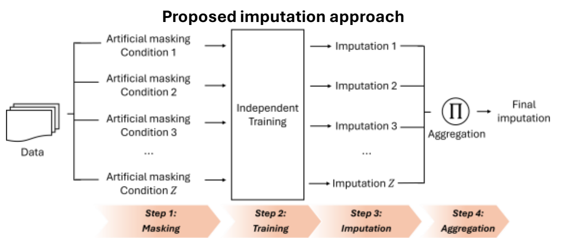
Time	Sensor 358477428	Sensor 916022554	Sensor 358423364
01/05/2023 00:00	18	23	19
01/05/2023 01:00	13	11	15
01/05/2023 02:00	6	9	6
01/05/2023 03:00	6	?	5
01/05/2023 04:00	7	13	7
01/05/2023 05:00	11	14	10
01/05/2023 06:00	?	50	19
01/05/2023 07:00	?	50	35
01/05/2023 08:00	65	51	?
01/05/2023 09:00	104	53	106



12

12

Data imputation (ensemble of imputers)

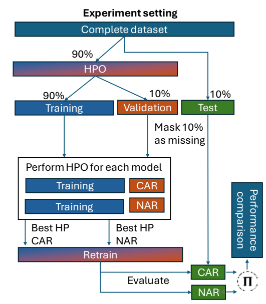


13

13

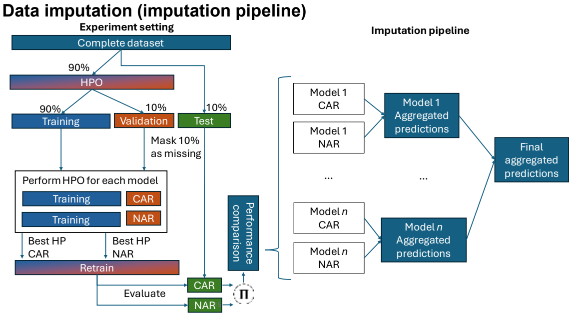
Data imputation (validation)

- Neural network-based
  - Self-Attention-based Imputation for Time Series (SAITS)
  - Bidirectional Recurrent Imputation for Time Series (BRITS)
  - Transformer-based imputation (Transformer)
  - Unsupervised GAN for Multivariate Time Series Imputation (USGAN)
- Naïve:
  - Last Observation Carried Forward (LOCF)
- General ML algorithms
  - IterativeImputer
  - KNNImputer
  - Simple imputers
- Matrix completion-based
  - Iterative soft thresholding of SVD decompositions (SoftImpute)
  - Iterative low-rank SVD decomposition (IterativeSVD)
  - Direct factorization of the incomplete matrix into low-rank U and V (MatrixFactorization)



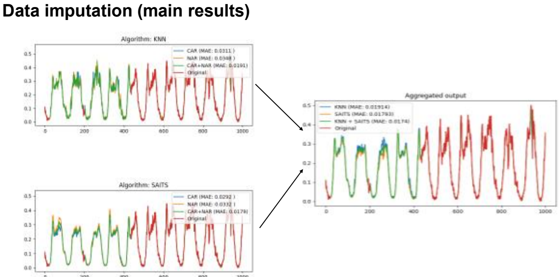
14

Data imputation (imputation pipeline)



15

Data imputation (main results)

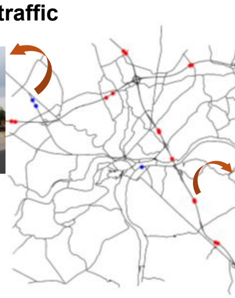


16

16

Data fusion in traffic

ANPR cameras



Counting loops

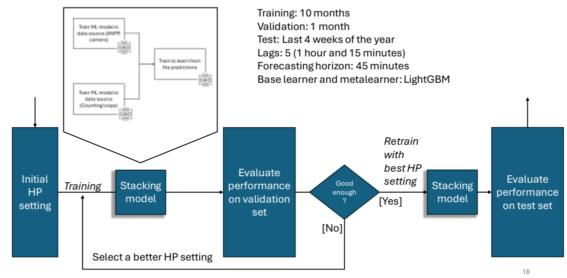
Data collected in 2023

17

17

Data fusion (ensemble of forecasters)

Training: 10 months
Validation: 1 month
Test: Last 4 weeks of the year
Lags: 5 (1 hour and 15 minutes)
Forecasting horizon: 45 minutes
Base learner and metalearner: LightGBM



18

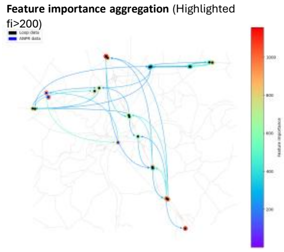
18

Data fusion (main results)

Mean Absolute Error

	Validation set		Test set	
	ANPR source	**Counting loop source**	**ANPR source**	**Counting loop source**
Stacking approach	3.3994	8.0415	2.9773	8.4754
Individual training (before using the metalearner)	3.8994	28.4025	3.382	27.4506
LightGBM trained without distinguishing between data sources	3.4585	8.8612	2.9813	9.1121

Feature importance aggregation (Highlighted fi>200)

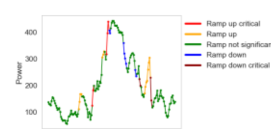


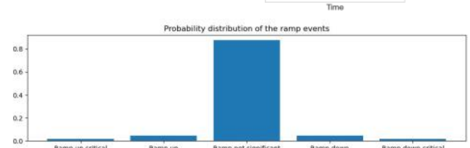
19

19

Ramp event forecasting

- A ramp event represents a large and fast variation in power in a wind farm or portfolio
- Most of the approaches are threshold-based methods for binary classification
- Unbalanced classification

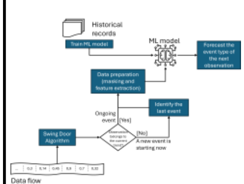




20

20

Ramp event forecasting



Algorithm 1 *EasyEnsemble*

Require: $\mathcal{P}$: A set of non-majority class examples
Require: $\mathcal{M}$: A set of majority class examples
Require: $\mathcal{L}$: The number of subsets $\mathcal{L}$ to sample from M
Require: S: The number of weak classifiers in an AdaBoost ensemble $\mathcal{H}$

1: **while** $i = 1, \ldots, \mathcal{L}$ **do**
2: Randomly sample a subset N_i from $\mathcal{M}$, $|\mathcal{N}_i| \approx |\mathcal{P}|$
3: Learn $\mathcal{H}_i$ using $\mathcal{P}$ and $\mathcal{N}_i$. $\mathcal{H}_i$ is an AdaBoost ensemble with S weak classifiers $h_{i,j}$ and corresponding weights $\alpha_{i,j}, j = 1, \ldots, S$. The ensemble's threshold is θ_i,

$$\mathcal{F}_i = sgn\left(\sum_{j=1}^{S} \alpha_{i,j} h_{i,j}(x) - \theta_i\right)$$

4: **end while**
5: **return** an ensemble

$$\mathcal{F} = sgn\left(\sum_{i=1}^{L}\sum_{j=1}^{S} \alpha_{i,j} h_{i,j}(x) - \sum_{i=1}^{L} \theta_i\right)$$

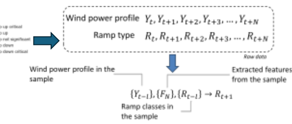
* X. Y. Liu, J. Wu and Z. H. Zhou, "Exploratory Undersampling for Class-Imbalance Learning," in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 39, no. 2, pp. 539-550, April 2009.

21

21

Ramp event forecasting

Proposed approach



Main results:

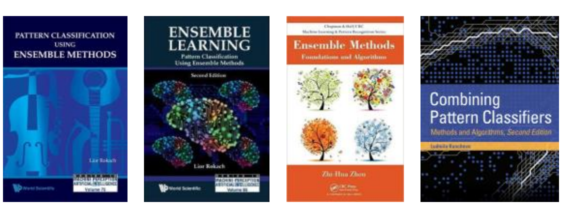
- Increasing *lags* did not offer any improvement
- The probabilistic classifier is a strong baseline defined for the problem
- *EasyEnsemble* was the best classifier when reducing the classes and masking unknown events

	Accuracy	bacc	Weighted F1 Score	Accuracy	bacc	Weighted F1 Score
MBalancedBaggingClassifier	0.8516	0.6771	0.8691	0.8710	0.4869	0.8788
MBalancedRF	0.8760	0.3333	0.8182	0.0161	0.2000	0.0005
MRandomClassifier	0.3410	0.3397	0.4476	0.2458	0.2115	0.3516
MRandomForest	0.7717	0.7207	0.8124	0.7804	0.4775	0.8185
MRUSBoostClassifier	0.8354	0.5941	0.8494	0.8076	0.3596	0.8224
MEasyEnsembleClassifier	0.9123	0.9286	0.9214	0.8719	0.4546	0.8759
MProbabilisticClassifier						
	YES			NO		
	Reduce number of classes					

22

22

Interesting reads



23

23

Ensemble Learning: Building Smarter Models through Collaboration

24-11-2025

Alejandro Morales Hernández
Machine Learning Group
Université Libre de Bruxelles, Belgium

* Supported by INNOVIRIS through the TORRES project
https://torresml6.wordpress.com/

24